MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 724 August, 1983

THE SMOOTHEST VELOCITY FIELD AND TOKEN MATCHING SCHEMES

A.L. Yuille

ABSTRACT: This paper presents some mathematical results concerning the measurement of motion of contours. A fundamental problem of motion measurement in general is that the velocity field is not determined uniquely from the changing intensity patterns. Recently Hildreth & Ullman have studied a solution to this problem based on an Extremum Principle [Hildreth (1983), Ullman & Hildreth (1983)]. That is, they formulate the measurement of motion as the computation of the smoothest velocity field consistent with the changing contour. We analyse this Extremum principle and prove that it is closely related to a matching scheme for motion measurement which matches points on the moving contour that have similar tangent vectors. We then derive necessary and sufficient conditions for the principle to yield the correct velocity field. These results have possible implications for the design of computer vision systems, and for the study of human vision.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial research is provided in part by the Advanced Research Projects Agency of the Department of Defence under Office of Naval Research contract N00014-80-C-0505,

1. Introduction

The measurement and interpretation of visual motion plays an important part in artificial and biological vision systems. Measurement of motion from the more primitive data provided by photoreceptors or sensors is a suprisingly difficult computational problem, which has attracted much attention in recent years [for example, Fennema & Thompson (1979), Ullman (1979), Horn & Schunk (1981), Hildreth (1983), Nagel (1983); also see reviews in Thompson & Barnard (1981), Ullman (1981)]. This measurement can take place at different stages in the analysis of an image, for example, at the levels of (a) raw intensity or (b) intensity changes or edges. Marr and Ullman [Marr & Ullman (1981)] suggested that initial motion measurement in the human visual system takes place at the locations of zero-crossings of the image filtered with the laplacian of a gaussian. This provides strong motivation for studying motion measurements of moving contours. The motion of contours is also of interest in its own right. Direct comparisons can be made between the predictions of a proposed computation and psychophysical results indicating the perceived motion. In this paper we restrict ourselves to such motion.

As a consequence of the *local* nature of the initial motion measurement only one component of velocity can be obtained directly from the changing image [Fennema & Thompson (1979), Horn & Schunk (1981)] (it may be possible, however, to obtain both components at places where the gradient of the intensity is discontinuous). Marr and Ullman referred to this as the aperture problem [Marr & Ullman (1981)]. In the case of moving contours only the component of velocity normal to the contours can be measured directly; the tangential component of the velocity field is undetermined. The computation of the full two-dimensional velocity field requires the integration of local motion measurements, either along the contour or over areas of the image. A fundermental theoretical problem for this integration stage is that for the case of general motion, the velocity field is not determined uniquely from the changing image. Additional constraints are required to compute a unique velocity field.

Recently Hildreth and Ullman [Hildreth (1983), Ullman & Hildreth(1983)] have proposed a solution to the motion measurement problem based on an Extremum Principle. For a contour with arc length parameter s they suggest selecting the velocity field $\underline{v}(s)$ which has the correct normal component and which minimizes the integral I along the contour

$$I = \int \left(\frac{\partial \underline{v}}{\partial s} \cdot \frac{\partial \underline{v}}{\partial s}\right) ds. \tag{1.1}$$

This corresponds to choosing the smoothest velocity field consistent with the

changing image. We will refer to this as the smoothness principle. The use of smoothness to constrain the velocity field is based on the physical assumption that surfaces are generally smooth, compared with their distance from the viewer, and therefore generate smoothly varying velocity fields when they move. A formulation of the smoothness constraint in two-dimensions had previously been proposed by Horn and Schunck [Horn & Schunck (1981)].

In this paper we derive some mathematical consequences of this principle. In section (2) we show that there is a close connection between this type of velocity field computation and one that establishes an explicit correspondence between points on the contour at different times, which have similar tangent vectors. Thus there is a close similarity between the smoothness principle and "token matching schemes" for motion measurement [Potter (1977) ,Ullman (1979), Thompson & Barnard (1981)]. This is somewhat surprising since the two approaches formulate the motion measurement problem in different, though roughly equivalent, ways. Rather than considering the instantaneous velocity field, the token matching approach takes "snapshots" of the contour at different times. The motion measurement problem then becomes: given two snapshots of a contour at different times, how do you match points on the two contours.

In section (3) we derive a necessary and sufficient condition for when this principle yields the correct result. While the computation of the correct velocity field may be important for the design of computer vision systems, it should be noted that the correct solution is not necessarily the one found by the human vision system. In situations where the smoothness principle does not yield the correct (true) velocity field, it appears that the smoothest velocity field may be more consistent with human motion perception. For example, Hildreth [Hildreth (1983)] shows that the perception of expansion or contraction when viewing a rotating spiral, the perceived downward motion of a rotating barberpole, and non-rigid appearance of a rotating ellipse are more consistent with the smoothest velocity field than the true one.

2. Matching Tangent Vectors

The approach based on the Extremum Principle (1.1) measures the normal component of the instantaneous velocity and computes a velocity field. Another approach [Ullman (1979)] to solving the motion measurement problem is, given two snapshots of the curve, to match points with similar tangent vectors. This essentially minimizes local deformation in the contour over time. If we denote the unit tangent vector by $\underline{T}$ and take the snapshots at an infinitesimal time apart, this is equivalent to minimizing a measure J where

$$J = \int \left(\frac{\partial \underline{T}}{\partial t} \cdot \frac{\partial \underline{T}}{\partial t}\right) ds. \tag{2.1}$$

We show that there is a close connection between this measure and the one proposed by Hildreth and Ullman and that they give similar predictions for a wide range of motions.

Let the image curve be denoted by

$$\underline{r} = \underline{r}(\theta, t). \tag{2.2}$$

Here t is the time and θ is a parametrization of the curve. The velocity distribution is

$$\underline{v}(\theta, t) = \frac{\partial \underline{r}}{\partial t} \tag{2.3}$$

where the partial derivative with respect to t is taken with θ constant. This means that the velocity distribution $\underline{v}(\theta, t)$ obtained depends on the parameter θ. If we choose another parameter $\psi(\theta, t)$ in place of θ we would get a different distribution. Solving the motion measurement problem corresponds to finding a preferred parametrization for the curve. This parametrization will show which points correspond in different snapshots and hence determines a velocity distribution.

If s is the arc length of the curve we calculate

$$\frac{\partial \underline{v}}{\partial s} = \left(\frac{\partial s}{\partial \theta}\right)^{-1} \frac{\partial \underline{v}}{\partial \theta}. \tag{2.4}$$

Using (2.3) we find

$$\frac{\partial \underline{v}}{\partial s} \cdot \frac{\partial \underline{v}}{\partial s} = \left(\frac{\partial s}{\partial \theta}\right)^{-2} \frac{\partial^2 \underline{r}}{\partial \theta \partial t} \cdot \frac{\partial^2 \underline{r}}{\partial \theta \partial t}. \tag{2.5}$$

Thus the smoothness principle is equivalent to finding the parameter θ which minimizes

$$\int \left(\frac{\partial s}{\partial \theta}\right)^{-2} \frac{\partial^2 \underline{r}}{\partial \theta \partial t} \cdot \frac{\partial^2 \underline{r}}{\partial \theta \partial t} ds \tag{2.6}$$

subject to the observed normal velocity.

The unit tangent vector is

$$\underline{T} = \left(\frac{\partial s}{\partial \theta}\right)^{-1} \frac{\partial \underline{r}}{\partial \theta}. \tag{2.7}$$

Differentiating with respect to t, we obtain

$$\frac{\partial \underline{T}}{\partial t} = -\left(\frac{\partial s}{\partial \theta}\right)^{-2} \frac{\partial^2 s}{\partial t \partial \theta} \frac{\partial \underline{r}}{\partial \theta} + \left(\frac{\partial s}{\partial \theta}\right)^{-1} \frac{\partial^2 \underline{r}}{\partial \theta \partial t}. \tag{2.8}$$

Hence

$$\frac{\partial \underline{T}}{\partial t}\cdot\frac{\partial \underline{T}}{\partial t}=\left(\frac{\partial s}{\partial \theta}\right)^{-4}\left(\frac{\partial^2 s}{\partial t\partial \theta}\right)^{2}\frac{\partial \underline{r}}{\partial \theta}\cdot\frac{\partial \underline{r}}{\partial \theta}+\left(\frac{\partial s}{\partial \theta}\right)^{-2}\frac{\partial^2 \underline{r}}{\partial \theta\partial t}\cdot\frac{\partial^2 \underline{r}}{\partial \theta\partial t}-2\left(\frac{\partial s}{\partial \theta}\right)^{-3}\frac{\partial^2 s}{\partial t\partial \theta}\frac{\partial^2 \underline{r}}{\partial \theta\partial t}\cdot\frac{\partial \underline{r}}{\partial \theta}. \tag{2.9}$$

Now by definition of s we have

$$\left(\frac{\partial s}{\partial \theta}\right)^{2}=\frac{\partial \underline{r}}{\partial \theta}\cdot\frac{\partial \underline{r}}{\partial \theta}. \tag{2.10}$$

Differentiating with respect to time we obtain

$$\frac{\partial^2 \underline{r}}{\partial \theta\partial t}\cdot\frac{\partial \underline{r}}{\partial \theta}=\frac{\partial s}{\partial \theta}\left(\frac{\partial^2 s}{\partial t\partial \theta}\right). \tag{2.11}$$

Substituting (2.10) and (2.11) into (2.9) gives

$$\frac{\partial \underline{T}}{\partial t}\cdot\frac{\partial \underline{T}}{\partial t}=\left(\frac{\partial s}{\partial \theta}\right)^{-2}\frac{\partial^2 \underline{r}}{\partial \theta\partial t}\cdot\frac{\partial^2 \underline{r}}{\partial \theta\partial t}-\left(\frac{\partial s}{\partial \theta}\right)^{-2}\left(\frac{\partial^2 s}{\partial t\partial \theta}\right)^{2}. \tag{2.12}$$

We now combine (2.12) and (2.5) and obtain

$$\frac{\partial \underline{v}}{\partial s}\cdot\frac{\partial \underline{v}}{\partial s}=\frac{\partial \underline{T}}{\partial t}\cdot\frac{\partial \underline{T}}{\partial t}+\left(\frac{\partial}{\partial t}log\frac{\partial s}{\partial \theta}\right)^{2}, \tag{2.13}$$

so we see that the smoothness measure can be written as the sum of two terms both of which are positive definite. The first term measures the change in the local tangent vectors at corresponding points on the contour over time, and can therefore be minimized by matching points with similar tangent vectors. The second term is a local measure of how fast the curve is expanding.

Since the right hand side of (2.13) is the sum of two positive terms it will be small only when both of them are small. The two terms are functionally dependent, so minimizing their sum is *not* equivalent to minimizing both of them individually. The second term becomes important in extended areas of the contour with zero curvature (here the correspondence between points with similar tangent vectots is ambiguous), and situations where the contour is expanding and distorting significantly. Note that the measures I and J are integral measures; the minimization of the integral is not equivalent to the minimization of local changes independently.

The result (2.13) is surprising since it relates two apparently different ways of modeling the motion of a contour. Note that if the contour is a straight rigid rod moving in the image plane the tangent measure will be zero for all possible motions and so will not determine a unique motion. We will show in the next section that the presence of the expansion term means that the smoothness measure will yield the correct result in this case, provided the

boundary conditions at the ends of the rod are known. Further investigation is needed to find the conditions under which the expansion term in the equation can be neglected. Examples for which this term is significant may determine which formulation is more appropriate for human vision.

Equation (2.13) also suggests that some of the nice results of token matching schemes will also apply to the Extremum Principle (1.1), and can be incorporated into a motion measurement computation based on it. Suppose, for example, that a contour has cusps, discontinuities or other easily distinguished points. Most token matching schemes will match these points between snapshots [Lawton (1983)]. For example, if the contour is a square the distinguished points are the vertices and they are matched between snapshots. In particular the tangent measure (2.1) will tend to match such points since they occur at places where the derivative along the curve of the tangent vector is a maximum (i.e. the modulus of the curvature is a maximum). If this maximum is infinite (i.e. if the contour has a discontinuity) such points must be matched since otherwise the integral (2.1) will become infinite. This behaviour is confirmed by experiment and, by (2.13), we expect it to be predicted by the smoothness principle.

3. The Validity of the Smoothness Measure

To test the validity of the smoothness principle for motion measurement, we can examine the predicted solution for actual moving contours. One method for obtaining this solution is to specify an algorithm that embodies this principle, and run the algorithm on a variety of moving contours [Hildreth (1983)]. A second, more direct method is to apply the Calculus of Variations to (1.1) and obtain the Euler-Lagrange equations for $\underline{v}$. The solution of these equations yields the correct velocity distribution but, as we will show, these equations are too complicated to be solved in general. There is an alternative way in which the equations can be used however; we can substitute the correct velocity distribution and determine whether it is a solution. Thus we can obtain a simple condition for when the smoothness principle yields the correct velocity field solution.

Let the curve in the image plane be $\underline{r} = \underline{r}(s)$ and the correct velocity be $\underline{X}$. Note that this is not the same as the expression for $\underline{r}$ in the previous section. The component of velocity normal to the curve, which we denote by $\underline{U}$, is

$$\underline{U} = (\underline{X} \cdot \underline{N})\underline{N} \tag{3.1}$$

where $\underline{N}$ is the unit normal to the curve. This component of the velocity is measured directly from the changing image, but the tangential component is unknown. Denote it by $F(s)\underline{T}$ where F is an arbitrary function. Then the full velocity $\underline{V}$ is written

$$\underline{v} = (\underline{X} \cdot \underline{N})\underline{N} + F(s)\underline{T}. \tag{3.2}$$

Before proceeding we recall some results from differential geometry [Faux & Pratt (1979)]. The curve lies in the image plane and hence has no torsion. The derivatives with respect to the arc length s of the tangent and normal unit vectors are given by

$$\frac{\partial T}{\partial s} = \kappa \underline{N} \tag{3.3}$$

and

$$\frac{\partial N}{\partial s} = -\kappa \underline{T} \tag{3.4}$$

where κ is the curvature.

We now return to (3.2). Differentiating with respect to s and using (3.3) and (3.4) we obtain

$$\frac{\partial \underline{v}}{\partial s} = \left(\frac{\partial \underline{X}}{\partial s} \cdot \underline{N} - \kappa(\underline{X} \cdot \underline{T}) + F\kappa\right)\underline{N} + \left(\frac{\partial F}{\partial s} - \kappa(\underline{X} \cdot \underline{N})\right)\underline{T}. \tag{3.5}$$

We substitute this into the smoothness measure to obtain

$$\int \frac{\partial \underline{v}}{\partial s} \cdot \frac{\partial \underline{v}}{\partial s} ds = \int \left(\left(\frac{\partial \underline{X}}{\partial s} \cdot \underline{N} - \kappa(\underline{X} \cdot \underline{T}) + F\kappa\right)^2 + \left(\frac{\partial F}{\partial s} - \kappa(\underline{X} \cdot \underline{N})\right)^2\right) ds. \tag{3.6}$$

We write the integrand of (3.6) as $L(F, \frac{\partial F}{\partial s})$, a function of F and $\frac{\partial F}{\partial s}$. By standard results of the Calculus of Variations [Courant & Hilbert (1953)] extremizing the integral (3.6) is equivalent to solving the Euler-Lagrange equations

$$\frac{\partial L}{\partial F} = \frac{\partial}{\partial s}\left(\frac{\partial L}{\partial F_s}\right) \tag{3.7}$$

where we use F_s to denote $\frac{\partial F}{\partial s}$. We have

$$\frac{\partial L}{\partial F} = 2\kappa\left(\frac{\partial \underline{X}}{\partial s} \cdot \underline{N} - \kappa(\underline{X} \cdot \underline{T}) + F\kappa\right) \tag{3.8}$$

and

$$\frac{\partial L}{\partial F_s} = 2\left(\frac{\partial F}{\partial s} - \kappa(\underline{X} \cdot \underline{N})\right). \tag{3.9}$$

We differentiate to obtain

$$\frac{\partial}{\partial s}\left(\frac{\partial L}{\partial F_s}\right) = 2\frac{\partial^2 F}{\partial s^2} - 2\frac{\partial \kappa}{\partial s}(\underline{X}\cdot\underline{N}) - 2\kappa\frac{\partial \underline{X}}{\partial s}\cdot\underline{N} + 2\kappa^2(\underline{X}\cdot\underline{T}). \quad (3.10)$$

Substituting (3.8) and (3.10) into (3.7) gives

$$\frac{\partial^2 F}{\partial s^2} - \frac{\partial \kappa}{\partial s}(\underline{X}\cdot\underline{N}) - 2\kappa\frac{\partial \underline{X}}{\partial s}\cdot\underline{N} + 2\kappa^2(\underline{X}\cdot\underline{T}) - F\kappa^2 = 0. \quad (3.11)$$

This equation is complicated and it is unlikely that it can be solved in general. Notice, however, that if the image curve is made up of straight lines with κ zero (for example if it is a square) the equation reduces to

$$\frac{\partial^2 F}{\partial s^2} = 0 \quad (3.12)$$

and can be solved.

Rather than looking for solutions of (3.11) we now ask, under what conditions do the equations yield the correct solution? (As mentioned in the introduction it should be emphasized that the human visual system does not always obtain the correct result). This occurs when the solution of (3.11) is

$$F = (\underline{X}\cdot\underline{T}). \quad (3.13)$$

Using formulae (3.3) and (3.4) we find

$$\frac{\partial F}{\partial s} = \left(\frac{\partial \underline{X}}{\partial s}\cdot\underline{T}\right) + \kappa(\underline{X}\cdot\underline{N}) \quad (3.14)$$

and

$$\frac{\partial^2 F}{\partial s^2} = \frac{\partial \kappa}{\partial s}(\underline{X}\cdot\underline{N}) + 2\kappa\left(\frac{\partial \underline{X}}{\partial s}\cdot\underline{N}\right) - \kappa^2(\underline{X}\cdot\underline{T}) + \frac{\partial^2 \underline{X}}{\partial s^2}\cdot\underline{T}. \quad (3.15)$$

Substituting (3.13) and (3.15) into (3.11) we find it reduces to

$$\frac{\partial^2 \underline{X}}{\partial s^2}\cdot\underline{T} = 0. \quad (3.16)$$

Thus (3.16) is a necessary and sufficent condition for the measure (1.1) to yield the correct solution. If the velocity field corresponds to uniform translation or expansion it is clear this condition is satisfied.

Suppose the contour rotates. The velocity field can be expressed as

$$\underline{X} = \underline{\omega}\times\underline{r} \quad (3.17)$$

where $\underline{\omega}$ is perpendicular to the image plane. Now

$$\frac{\partial^2 \underline{X}}{\partial s^2} = \kappa \underline{\omega} \times \underline{N} \tag{3.18}$$

and substituting this into (3.16) gives

$$\kappa(\underline{\omega} \times \underline{N}) \cdot \underline{T} = 0. \tag{3.19}$$

Since $\underline{\omega}$, $\underline{N}$ and $\underline{T}$ are mutually orthogonal, (3.19) holds if and only if κ is zero. Thus the smoothness principle only gives correct results for rigid rotation when the curve is made up of straight line segments (with κ zero). At the vertices κ has a discontinuity of delta function type. The boundary condition this imposes requires that the velocity field is continuous at the vertices.

In fact if the curve is made up of straight line segments, which do not deform, the smoothness principle always yields the correct result. To see this, recall that the velocity of a straight line segment can be decomposed into rotation, translation and expansion components. Equation (3.16) implies that the correct rotation and translation components are found and the boundary conditions at the ends of the straight line segments ensures the correct expansion component.

4. Conclusion

This paper derives two results about the smoothness principle. The first suggests a close connection between it and token matching schemes. The connection with the tangent measure (2.1) is made explicit in (2.13). The second characterizes the smoothness principle and obtains a necessary and sufficient condition (3.16) for the principle to provide the right answer.

These results suggest that algorithms based on the smoothness principle will give correct results, and hence be useful for computer vision systems, when (a) motion can be approximated locally by pure translation, rotation or expansion, or (b) objects consist of connected straight lines. In other situations, the smootness principle will not yield the correct velocity field, but may yield one that is qualitatively similar [Hildreth (1983)]. If the measurement of motion in the human visual system uses the smoothness principle, the results suggest when the human system should derive the correct solution. This could be tested by perceptual experiments. Finally different formulations of the principle may be useful for different purposes. Formulations like (1.1) led straightforwardly to a proof of the uniqueness of the result [Hildreth (1983)], but formulation in terms of tangent matching and local expansion may lead to a simpler algorithm.

Aknowledgements

I would like to thank Ellen Hildreth, Tommy Poggio and Shimon Ullman

for helpful discussions and Ellen Hildreth for many useful suggestions and comments on drafts of this paper.

References

Courant, R. & Hilbert, D. *Methods of Mathematical Physics,* Interscience Publishers, Inc, New York, 1953.

Faux, I.D. & Pratt, M.J.. *Computational Geometry for Design and Manufacture,* Ellis Horwood Limited, Chichester, England, 1979.

Fennema, C.I. & Thompson, W.B. "Velocity determination in scenes containing several moving objects," *Comp. Graph. Image. Processing* **9** (1979), 301-315.

Hildreth, E.C. "The computation of the velocity field," *M.I.T. A.I. Memo* **734** (1983), .

Horn, B.K.P. & Schunck, B.G. "Determining optical flow," *Artif. Intell.* **17** (1981), 185-203.

Lawton, D.T. "Processing translational motion sequences," *Comp. Vision. Graph. Image Processing* **22** (1983), 116-144.

Marr, D. & Ullman, S. "Directional selectivity and its use in early visual processing," *Proc. Roy. Soc. Lond.* B.204 (1981), 151-180.

Potter, J.L. "Scene segmentation using motion information," *Comp. Vision. Graph. Image Processing* **6** (1977), 558-581.

Nagel, H-H. "Constraints for the estimation of displacement vector fields from image sequences," *Proceedings IJCAI-83* (1983), .

Thompson, W.B. & Barnard, S.T. "Lower-level estimation and interpretation of visual motion," *IEEE Computer.* August (1981), 20-28.

Ullman, S. "Analysis of visual motion by biological and computer systems," *IEEE Computer* August (1981), 57-69.

Ullman, S. *The Interpretation of Visual Motion,* MIT Press, Cambridge, Mass., 1979.

Ullman, S & Hildreth, E.C. "The measurement of visual motion," *In Physical and biological processing of images.* ed. Braddick, O.J. & Sleigh, A.C. (1983), Berlin: Spinger-Verlag.